Explanatory notes on words marked * in the frequency lists

*ai* This is the initial part of *ain’t*. It has not been assigned to a lemma, because of its ambiguity: in different contexts, it can be regarded as a reduced form of *am, is, are, has, have*, etc.

*all right, alright* These variant spellings have been listed and counted separately.

*be* The count for the lemma *be* is slightly underrepresented, because of the omission of *ain’t* and *in~* (see *ai~* above and *in~* below).

*because* is listed separately from the shortened spoken forms *cos* and ’*cos* (q.v.).

*bit* Note that the count for *bit* as a noun excludes tokens of *a bit* as a multiword adverb.

*conservative(s), labour, liberal(s)* Whether as adjectives or nouns, these counts include references to political parties (usually spelt with a capital), as well as more general senses. These words are not capitalized in the lists.

*cos*, ’*cos* These shortened forms of *because* are given separate entries in the lists.

*course* (Adv), as a shortened form of *of course*, is given a separate entry.

*datum* It is no longer realistic to treat *data* as the plural of *datum*, because most people these days treat *data* as a singular (uncountable) noun. Hence, the lemma is given the combined label *data/datum*.

*de* When *de* is part of a foreign name (e.g. *de Gaulle*), it is counted as a NoP-. Elsewhere, e.g. in French quotations, it is counted as a foreign word (Fore).

*du~* This is the first part of *dunno* (= *do + not + know*), and is lemmatized with *do*.

*elite* The counts of *elite* (NoC) do not include tokens with the acute accent (*élite NoC*), of which there are about 3 per million words in the whole corpus.

*fig* This word is ambiguous between *fig* (a fruit) and *fig* (also spelt *figg*) as the abbreviation for *figure*. The two usages are not distinguished here; *fig* (’fruit”) is the rarer usage, but is more widely scattered through the whole corpus.

*further* (Adj): As an adjective, *further* is lemmatized as a comparative form of *far*. However, in meaning and function, *further* is not always comparative: it often means ‘additional’ and ‘extra’.

*goin’, going* as forms of the verb *go*, exclude the semi-auxiliary verb *be going to* (see below).

*going (to)* This shows the frequency of the semi-auxiliary verb *be going to*, which is given a separate count, rather than being included as a form of *go*.

*good* (NoC): This is here treated as the lemma to which the plural form *goods* belongs. Note, however, that *goods* (in the sense of ’freight’) has no corresponding singular form.

*have* The count for the lemma *have* is slightly underrepresented, because of the omission of the form *ain’t* (see *ai~* above).
her  (Det): As a possessive determiner (as in her friends), her is treated as a separate lemma from her as a personal pronoun (in the objective case).

his  (Pron): As a possessive pronoun (as in This is his), his is treated as a separate lemma from his as a possessive determiner (as in It was his fault).

in~  (Verb): The first part of the vernacular British tag question innit. Like ai~, it is difficult to assign this verb to be, have or any other verb, because it is commonly used across a range of functions.

Jan  (NoP): This entry is ambiguous between an abbreviation for January (also written Jan.) and a person’s given name.

labour  see conservative(s) above.

lead  (NoC): Notice that this headword is ambiguous, according to whether it is pronounced as /lɛd/ or as /lɛd/. Both pronunciations are included in the count for lead.

Les  (NoP): Counted as a proper noun, Les can be either a given name (= Leslie) or a part of a French name, as in Les Routiers.

liberal(s)  see conservative(s) above.

lot  (NoC): Note that the count for lot as a noun excludes tokens of a lot as an adverb.

me  (Det): As a possessive determiner, my sometimes occurs in the non-standard form me (as in I’ll ask me dad). Notice that this form is distinct from me as an objective pronoun (e.g. Don’t ask me).

miss  (NoC): This lemma is ambiguous, and includes both Miss/miss as a title of address for an unmarried woman, and miss in the sense of ‘a near miss’.

more than  (Adv): More than is considered to be a multiword adverb of degree when it is followed by a numerical expression, as in It costs more than £100 (= ‘over’).

~n~  A variant of the negative not: the middle part of dunno (= do + not + know).

need  (VMod): This count includes only the tokens of need as a modal auxiliary: e.g. where need is followed by a negative and/or a bare infinitive: need not go, needn’t leave, I doubt if you need say anything. This is counted separately from need (main verb).

~na  This is a rendering of the infinitive marker to, when run together with the preceding verb in gonna and wanna.

~no  The final part of dunno (= do + not + know), lemmatized with know.

no one, no-one  The two variant spellings of this pronoun appear in separate entries.

of  (Prep): Note that the count for of as a preposition excludes the many cases of multi-word prepositions or adverbs including of: e.g. of course, in spite of.

of  (Verb): In the transcriptions of speech for the BNC, a weakened pronunciation of have was sometimes written as of (as in could of done it).

okay, OK  Note that okay, classified as an adverb or an adjective, also has a variant spelling OK, with a separate entry.

~ta  This is an informal spelling of the infinitive marker to, especially in the combination gotta (= got+ to).